DOCUMENT RESUME

ED 158 861 PS 009 564

AUTHOR Tharp, Roland G.; Gallimore, Ronald
TITLE The Uses and Limits of Social Reinforcement and Industriousness for Learning to Read. Technical Report #60.
INSTITUTION Kamehameha Schools, Honolulu, Hawaii. Kamehameha Early Education Project.
SPONS AGENCY California Univ., Los Angeles. Mental Retardation Research Center.; Hawaii State Dept. of Education, Honolulu.
PUB DATE [76]
NOTE 32p.; For related documents, see PS 009 533-563 and PS 009 565-573

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Attention Control; Curriculum Development; *Demonstration Programs; Early Childhood Education; Educational Improvement; Hawaiians; Public Schools; *Reading Achievement; *Reading Improvement; Research; *Social Reinforcement; Study Habits; *Teacher Behavior
IDENTIFIERS Hawaii; *Kamehameha Early Education Program

ABSTRACT

This report describes a study of the use of social reinforcement to increase the industriousness, and subsequently the reading competence, of children in the kindergarten through third grade classes of the Kamehameha Early Education Program (KEEP) demonstration school. Teacher behaviors, pupil industriousness, and pupil reading performance were measured in the KEEP school and in classes in five Hawaiian public schools. Comparisons between these measures demonstrated that KEEP teachers use more praise than teachers in the public school classes and that KEEP children are more industrious than public school children. There were no concomitant differences, however, in reading achievement scores. It is concluded that, while industriousness is important, it is not a sufficient prerequisite for the improvement of reading skills. It is suggested that improvement in reading achievement will depend on the development of a curriculum appropriate for the Hawaiian and part-Hawaiian populations represented in KEEP and in the Hawaiian public schools. (Author/BD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

- ED158861

Technical Reports

of

The Kamehameha Early Education Program

a research and development program established and funded by

The Kamehameha Schools/Bernice P. Bishop Estate

Ronald Gallimore, Roland G. Tharp & Gisela E. Speidel,
General Editors

Technical Report #60

Ellen Antill
Production Editor

The cooperation of the State of Hawaii Department of Education is gratefully acknowledged, as is the support and resources made available by the Sociobehavioral Research Group, MRRC, University of California, Los Angeles.

PS009564

The opinions expressed herein do not necessarily reflect the position, policy or have the endorsement of The Kamehameha Schools/Bernice P. Bishop Estate, or of the editors.

Published by The Kamehameha Early Education Project, 1850 Makuakane Street, Honolulu, HI 96817

All rights reserved. No parts of this report may be reproduced in any form or by any means without the prior written permission of The Kamehameha Schools/Bernice P. Bishop Estate.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kim C. M. Sloat

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

The Kamehameha Early Education Program

The Kamehameha Early Education Program (KEEP) is a research and development program of The Kamehameha Schools/Bernice P. Bishop Estate. The mission of KEEP is the development, demonstration, and dissemination of methods for improving the education of Hawaiian and Part-Hawaiian children. These activities are conducted at the Ka Na'i Pono Research and Demonstration School, and in public classrooms in cooperation with the State Department of Education. KEEP projects and activities involve many aspects of the educational process, including teacher training, curriculum development, and child motivation, language, and cognition. More detailed descriptions of KEEP's history and operations are presented in Technical Reports #1-4.

Abstract

This report reviews KEEP's initial research strategy which was designed to test whether teachers' use of contingent social reinforcement would increase industriousness and thus reading competence in KEEP children, when curricular variables are held constant. Comparisons to control groups demonstrated that KEEP teachers do praise more and KEEP children are more industrious. Issues of causality and external validity are discussed. Although contingent social reinforcement significantly increased student industriousness, reading achievement was still below grade level. It is concluded that, while industriousness is important, achievement will depend on the development of a curriculum appropriate for populations similar to KEEP.

Technical Report #60

The Uses and Limits of Social Reinforcement and Industriousness for Learning to Read

Roland G. Tharp Ronald Gallimore

The Problem

Too often, by the end of third grade, Hawaiian children in the public school are already on the road toward school alienation. The destructive and ramifying effects of a failure to "keep up" in reading have been discussed often. For example, the KEEP children perform at or above national norms in early mathematics, but falter when they encounter word problems which they cannot read and comprehend. As with other disadvantaged populations, boredom, discontent, and frustration follow. Therefore, KEEP's primary goal is the analysis and solution of the problem of learning to read. This goal is consistent with that of the State Department of Education, and was even included as a first priority in the 1976 Governor's State of the State address to the Legislature, an indication of the extent of this problem.

The Strategy

Why are these children slow to learn reading? If anything, Hawaii's educators are given too many notions, and too little hard research evidence. This is particularly true of the population in question. Among the more promising lines of research that we are following are linguistic development, cognitive style, and cultural teaching styles. The line to be reported here involves motivation and social influence.

Extensive preliminary interviews of Island educators yielded many observations and opinions about Hawaiian low achievement, but the most frequent diagnosis was that of lack of motivation to learn. The children were regularly

described as "disinterested," "lazy," and "don't care." This pattern is familiar, and is a syndrome which has often responded to systematic programs of contingent reinforcement.

Our initial overall research strategy was designed to test whether contingent social reinforcement could be employed to increase reading competence in this population, by means potentially available to public school teachers, and manageable within realistic budgetary and taxation limits.

However, motivation is not the only reasonable hypothesis for reading deficiencies. Perhaps the reading curriculum must be adapted to meet the Hawaiian child's way of learning. For example, the vast majority of Hawaiian children speak Hawaiian Islands Creole English ("pidgin"), although the language of instruction is Standard English; perhaps linguistic issues are central. Also, the Hawaiian child, even though he may live in a modern urban environment, is part of a culture still heavily influenced by traditional patterns of social organization, traditional family teaching styles, and, to some degree, non-Western cognitive styles.

Had we attempted to manipulate all these variables simultaneously, a full understanding of effects would have been sacrificed. Our strategy called for an inquiry into motivational issues first, while "controlling" other variables to the degree possible in an actual school.[1] That is, while receiving contingent reinforcement for diligent work at the learning task, would our experimental subjects learn to read better than a comparable group in the public schools who used the same curriculum and language? The overall problem is conceptualized in Figure 1. Will a teacher who uses much contingent reinforcement

1. There is much evidence to support the effectiveness of reinforcement but most comes from laboratory studies. Recently, both Campbell (1973) and Cronbach (1975) have argued in favor of externally valid tests of the application of principles derived from laboratory research. Although our data may not meet some standards of internal validity, we believe that work done in a real setting can contribute much to the study of reinforcement.

produce a child who is industrious? Will this industriousness, given a standard curriculum, lead to reading? The simple form of this question is: If children can be taught through reinforcement to perform the behaviors which the school requests, will they learn to read?

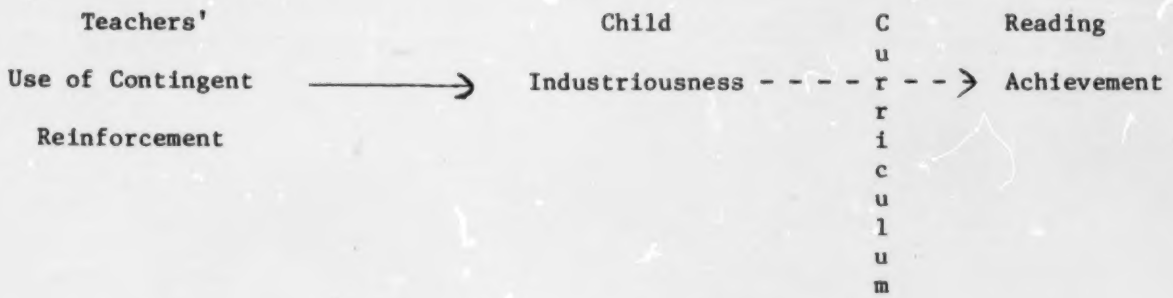

Figure 1

The answer to that question is not self-evident, but it is crucial to the overall inquiry of our research. If the answer is yes, there is a simple remedy to the reading problem, and issues of linguistics and cognition become moot. If the answer is no, we must attend to the behaviors prescribed by the schools: For this population, at least the curriculum would then be judged at fault.

Assessment required a demonstration that, compared to a control group, trained teachers did indeed employ some different pattern of reinforcement; experimental children were indeed more industrious; and the curriculum and language variables were equivalent. Then the question can be answered: Is there a difference in learning to read?

Selection of Social Reinforcement

For a variety of reasons, we focused on applications of social reinforcement. First, earlier research showed that teacher praise consistently raised

Hawaiian students' level of industriousness in the classroom (Gallimore, Boggs, and Jordan, 1974; MacDonald and Gallimore, 1971). (Gallimore et al. attributed these effects to cultural factors, but added that favorable response to positive social reinforcement has been found in other populations.)

Second, generally, teachers in Hawaii are more amenable to social reinforcement techniques than to token reinforcement systems (MacDonald and Gallimore, 1971).

Third, social reinforcement is clearly more practical as it is available to any teacher and clearly costs less than material rewards.

KEEP has made limited use of various token reinforcement techniques (see, for example, Technical Reports #10, #12, and #49), usually for problem situations or particular curriculum elements. For example, a particularly disruptive child was given sessions in a "Staats box" to shape up his classroom behavior (see Technical Report #9).

Teacher's Use of Social Reinforcement

During the four year period covered by this report, a total of 14 teachers have taught at the KEEP school. All were trained in reinforcement techniques by senior KEEP staff; none were familiar with the procedures and theory prior to employment. The procedures of training have been reported elsewhere (see Technical Reports #7 and #41); in brief, they included didactic instruction, guided participation, modeling, and feedback. The training procedures have been demonstrated reliable, even in brief workshop formats for public school teachers (see Technical Reports #33 and #52).

Social reinforcement techniques have become fairly well institutionalized in the KEEP school; they are maintained by various forms of feedback and by the training that the KEEP teachers offer to other public school teacher-trainees through workshops and direct consultation. Institutionalization, however, can lead to inaccurate self-perception, and it is necessary, even after four years, to measure our teachers' behavior from time to time, both to provide feedback

for their own self-regulation, and to compare with control subjects.

Measurement of Teacher Behaviors

For different purposes, we employ variations of a nine-category teacher-behavior observation schedule (see Technical Report #19). In most cases, only some categories occur with enough frequency and variance to be analyzed: These are verbal positive--divided into "academic" and "management" praise depending on the type of child behavior to which it refers--and verbal negative (or scolds), which refer almost totally to management behaviors, since (by definition) we exclude verbal feedback that an academic response is incorrect (unless the teacher's tone is scolding, which occurs very infrequently). Nevertheless, this issue of "tone" has been difficult to conceptualize, and, during 1973, we did not achieve sufficient reliability of observation to place confidence in our data, so they are not reported (even though the comparisons, such as they were, are favorable to us).

We do investigate nonverbal positive and negative responses, but only limited analyses can be presented here. Nonverbal negative responses (hitting, sending from the room, etc.) almost never occurred in any schools observed. Nonverbal positive responses are of two general types, both of great interest; we do not report them, even though the comparisons are highly favorable to us, for the following reasons. Type 1 is composed of tokens, marks on charts, etc. These systems were rarely employed in comparison schools; at KEEP, their frequency varied enormously depending on whether or not a brief token system was or was not being used, and thus the rate is very unstable. Type 2 is composed of hugs, pats, and affectionate physical contact. Again, their frequency at KEEP is higher than in comparison schools, but we do not present the data because of the wide variation among our own teachers, some of whom almost never touch the children and some of whom are extremely demonstrative in physical ways.

Thus the presented categories are Verbal Positive (academic and management praise), and, for 1974 and 1975, Verbal Negative. Over the years, inter-observer reliabilities have varied, but have ranged from overall .81 in 1973-74 to overall .96 in 1975.

We believe observations of our own teachers to be valid: Observers stand unseen in an observation deck, equipped with a one-way mirror with full vision and an extensive ceiling microphone system. They work continuously in this fishbowl, and reactivity effects, though probably present, are minimized.

For control group teachers, we observe from within classrooms, and reactivity effects are no doubt present. These teachers are probably aware of the general nature of our categories; therefore reactivity effects would underestimate observed differences between these teachers and KEEP teachers.

Control subjects were selected at random with respect to probable praise rates. In fact, we have taken almost any subjects who would agree to be observed. The necessary quid pro quo arrangements with these teachers has limited both the number of locations available. However, they range in distance from near neighbors to neighbor island, in setting from rural to urban, in size of classroom from large to small, though all have largely Hawaiian pupils of the same ages as our own. It is known that praise rates by teachers vary by age of child (White, 1975); they also vary by time of year, by subject matter, etc. Therefore, when making comparisons, we have paired control group data with observations on KEEP teachers taken from the same day (or when those data were unavailable, the nearest adjacent day), during similar activities, and from the same age children. Table 1 presents the means of these comparison days.

All available comparisons are summarized here (see Table 1). They are drawn from the two year period following our own achievement of reasonable stability of operation, to the present time. During one year, we had daily

Table 1

Summary of Mean Teacher Verbal Approvals and Disapprovals per 15 Minutes
KEEP and Comparison Schools

Group	Grade Level	No. of Teachers	No. of Obser.	Academic Praise $\bar{X}$	Academic Praise SD	Management Praise $\bar{X}$	Management Praise SD	Total Praise (A + M)	Verbal Negative $\bar{X}$	Verbal Negative SD
KEEP	K-1	3	12	27.08**	14.89	17.33**	11.82	44.41**		
School 1	K-1	3	9	3.44	2.22	.44	.50	3.88		
KEEP	K-1	3	12	27.08**	14.89	17.33**	11.82	44.41**		
School 2	K-1	3	10	3.70	4.21	2.70	3.55	6.40		
KEEP	K-1	2	9	23.88*	15.55	14.88*	11.05	38.76**		
School 3	K	2	17	10.76	10.40	6.35	8.06	17.11		
KEEP	K-2	5	25	31.60**	14.26	26.20**	18.00	57.80**	0	0
School 4	K-6	13	12	4.67	4.58	.95	2.18	5.58	4.96	7.23
KEEP	2-3	4	11	12.47	4.37	2.99**	1.82	15.46	3.99	1.78
School 5	2-3	3		12.54	9.56	.54	1.21	13.08	6.63	5.66

* $p < .025$

** $p < .005$

60-7

observations on all teachers; during others, we had weekly observations; at the present time, we make spot-checks. The data reported here are reasonable equivalents to our total yearly means, although we have chosen to use the means of the day-against-day comparisons as the most precise procedure available. The "chosen" days are in fact those on which we could persuade some control teacher for some reason to allow observation.

Data are available from five schools.

School 1. Three female teachers taught 58 kindergarten and first grade children in a single, combined Hawaii English Program classroom. These teachers were participants in a workshop in positive reinforcement held by KEEP in May of 1974 (see Technical Report #33). The data reported here are three days of baseline observations collected at the end of April by a KEEP research assistant. All data were collected in the morning during language arts. Means and standard deviations for academic and management praise are presented in Table 1.

KEEP teachers had significantly higher rates of academic (t (19)=4.68, $p < .005$) and management praise (t (19)=4.25, $p < .005$).

School 2. Three female teachers taught 46 kindergarten-first grade children in a single classroom. The data for them are also baseline since these teachers also participated in the same workshop as the three teachers from School 1. A KEEP research assistant observed for four days at the end of April; two teachers were absent for one of those days. All observations were done in the morning during language arts.

To make the comparison equivalent in time and subject matter, data are presented for three KEEP teachers who instructed reading and language arts in the morning to equivalent grade levels. Data were taken from the same time period in April during which the research assistant observed at School 1. Data

are based on frequency per 15 minutes.

Means and standard deviations for the observed behaviors are presented in Table 1. Again, KEEP teachers had a significantly higher mean rate of academic (t (20)=4.78, $p < .005$) and management praise (t (20)=3.75, $p < .005$).

School 3. At School 3, two self-contained kindergarten classes, each taught by one female teacher, were observed. Data were collected on these two teachers for three days in the fall of 1973, three days in the winter (with the exception of T2 who was absent one day), and three days in the spring of 1974. These data were collected solely for the purpose of comparison with KEEP teachers. All data were collected in the morning. During the fall, kindergarten readiness activities were observed; the rest of the data were collected during language arts activities.

Data are presented for the two KEEP teachers who taught kindergarten. An attempt was made to make the data equivalent but this could not be fully accomplished. Two observations of T1 took place in the afternoon, during math, and not in the morning. The remainder of the observations were taken when T1 taught in the morning. Also, often when observers were in the field, no one was at KEEP to collect the data on times corresponding to the control, so that adjacent days were used as comparison. T2's pupils were entirely kindergarten; T1 taught KEEP's combined K-1 section.

Although School 3's teachers' rates are relatively high, KEEP teachers gave over double the mean frequency of academic and management praise (A: t (24)=2.57, $p < .025$; M: t (24)=2.25, $p < .025$). (See Table 1)

For Schools 4 and 5, the coding procedure and the method of summarizing the data were altered. Observers recorded behavior for 15 minutes at five-second intervals. Within each interval, any behavior which occurred within that period was scored, but multiple occurrences (for example, two praises in five

seconds) were scored only once. In practice, this meant that many behaviors were not recorded because they did occur more than once within the interval.

Data were reduced to mean number of intervals in which a behavior occurred.

School 4. School 4 is located in a rural community on a neighbor island; data are available for 13 teachers. Seven taught the lower elementary grades, and six the middle elementary grades (see Technical Report #52).

These data are compared with those gathered on the five KEEP teachers for whom data were available during the same semester (Spring, 1975). These data also served as follow-up data for the KEEP teachers, who had completed training during January.

Means and standard deviations for the observed behaviors are presented in Table 1. Compared to School 4's teachers, KEEP teachers give significantly more academic praise (t (16)=7.42, $p < .005$) and management praise (t (16)=5.71, $p < .005$).

School 5. A three-on-two team of teachers was observed at School 5, a nearby public elementary school. These teachers taught second and third grade children.

The KEEP data were gathered on the four teachers currently instructing second and third grade in the KEEP classrooms. The data represent language arts and math. The data were collected during the same month as the comparison data were taken.

Table 1 presents the means and standard deviations on positive and negative feedback rates for the four KEEP teachers and for the three teachers from School 5. Management praise occurs significantly more at KEEP than at this comparison school (t (20)=3.46, $p < .005$).

An overall examination of Table 1 indicates that KEEP teachers have a higher rate of verbal praise than comparison schools.

Mainland Sample. White (1975) summarized sixteen observational studies

(apparently in the greater New York area) to determine natural rates of teacher verbal approval and disapproval in the classroom. Teacher approvals were divided into instructional (similar to academic praise) and managerial (similar to management praise) categories. Obviously, this "comparison" is of limited value, since there is no certainty that the categories are indeed equivalent; futhermore, White has no kindergarten data, and KEEP's rates are highest for K. We present it only because White's are the only "normative" data known to us.

White has grouped her data by grade and KEEP data will also be presented by grade to facilitate comparison. Data are given by means per group of teacher per grade level. KEEP data were gathered in November, 1975 on all eight teachers currently instructing in the KEEP classrooms. These data are chosen for comparison because they are our most recent and because they are the least favorable to us. Second and third grade classes are combined at KEEP. Therefore, the same data appear for second and third grades in Table 2.

Tables 1 and 2 seem to justify the assumption that KEEP teachers do manage their classrooms differently than their public school peers. KEEP teachers had higher overall praise rates, especially for management praise, and a much lower rate of disapproval.

Although the following summary is useful only for gross comparison, it gives a general picture of how much KEEP differs overall from all its comparisons (see Table 3).

Inspection of the overall means shows that KEEP teachers give more than three times as much praise on the average as do teachers in comparison schools.

We can conclude, then, that our experimental teachers do in fact differ from their colleagues in sister-schools in that they express notably more approval to their students, and notably less disapproval. Our teachers do what we purport that they do.

Table 2 [2]

Rates of Teacher Verbal Reinforcement

per 15 Minutes by Grade: KEEP and NY Sample [3]

Grade	Group	No. of Teachers	Academic	Management	Total Praise	Verbal Negative
K	KEEP	2	19.95	6.80	26.75	2.95
	NY	-	nda*	nda*	nda*	nda*
1	KEEP	2	5.80	5.50	11.33	3.16
	NY	33	9.07	1.12	10.21	7.60
2	KEEP	4	12.47	2.99	15.46	3.99
	NY	3	18.15	.15	19.50	10.35
3	KEEP	4	12.47	2.99	15.46	3.99
	NY	10	5.55	.15	5.70	7.05
Total	KEEP	8	12.86	4.47	16.94	3.69
	NY	46	10.46	.63	11.09	8.04

* No data available

---

2. Data adapted from M. A. White. Natural Rates of Teacher Approval and Disapproval in the Classroom. Journal of Applied Behavior Analysis, 1975, 8 (4), 367-372.

3. White's data represent raw frequency; KEEP's data are coded so that any one category of behavior can be entered no more than once per interval. Therefore, KEEP frequencies (mean number of intervals in which behavior occurred) represent underestimates as compared to the White system.

Table 3

Overall Means Teacher Approvals & Disapprovals per 15 Minutes by Manner of Data Collection

Group	Grade Level	Academic		Management		Total Praise	Verbal Negative	
Mean Absolute Frequency		$\bar{X}$	SD	$\bar{X}$	SD	(A+M)	$\bar{X}$	SD
KEEP 73	K-1	26.22	10.56	16.34	3.53	42.56	nda*	
School 1	K-1	3.44	2.22	.44	.50	3.88	nda	
School 2	K-1	3.70	4.21	2.70	3.56	6.40	nda	
School 3	K	10.76	10.40	6.35	8.06	17.11	nda	
Overall KEEP		26.22	10.56	16.34	3.53	42.56	nda	
Overall Comparison		5.32	3.82	2.82	3.21	9.13	nda	
Mean Number Intervals								
KEEP Winter 74-75	K-2	31.60	14.26	26.20	18.00	57.80	0	0
KEEP Fall 75	K-3	12.86	8.78	4.47	2.84	17.33	3.69	3.44
School 4	K-6	4.67	4.58	.95	2.18	5.58	4.96	7.23
School 5	2-3	12.54	9.56	.54	1.21	13.08	6.63	5.66
NY	1-3	10.46	nda	.63	nda	11.09	8.04	nda
Overall KEEP		21.80	15.27	13.87	16.06	35.67	2.07	3.10
Overall Comparison		6.03	4.93	.65	.54	6.68	6.01	3.81

* no data available

But what is this pattern exactly? This line of operation and research has grown from the operant model, and begins with the assumption that teachers' positive social responses, when made contingent, will strengthen the children's behavior which they follow. In the real classroom setting, this is virtually impossible to document. Teachers who praise more must be reinforcing more, but the precise contingencies and schedules are lost in a morass of child behaviors. Even the teacher with the most delicate control of her own responses will reinforce by each praise a variety of behaviors, some desirable and some not, some known and others hidden. Over a long period of time and over many responses, the appropriate use of teacher praise should result in eventual discrimination and strenthening of the behaviors that the teachers value, but these behaviors are many and varied--from standing in line nicely to completing arithmetic work sheets to reading the story correctly to (apparently) listening to her. This overall response class, if there is one, is doing-what-the-teacher-wants, within the prescribed curriculum. It is what we call industriousness, and is measured by on-task rate.

Industriousness

By industriousness we mean behaving as prescribed by the teacher and the curriculum. In this sense, to be on-task is synonymous. During a school day, the prescribed behaviors vary: They may include attentiveness during instructions, cutting-and-pasting or drawing, moving directly from one learning center to another, or reading from a booklet; it can equally well include vigorous singing and dancing, or chasing moths for science bottles. The categorical definition of on-task is engagement in the behaviors appropriate to the teacher's instructions. We do not measure on-task during free play, so it never includes gossiping, dawdling, or idling.

On-task, as we measure it, is significantly related to actual work completed by the children; we correlated the average on-task and average number of work

pages completed during the same month (April, 1973) and obtained an r=.61 (see Technical Report #6).

In our own KEEP school, on-task observations are collected by the search method. Observers behind the one-way mirror search for each child in turn by name, and record his behavior as on-task, off-task, or disruptive. This has enabled us to keep longitudinal data for each child. Typically, three observations per day are made, with three consecutive searches per observation (about nine data points per day per child).

At control, or comparison, classrooms, we use the sweep technique because we do not know the children's names. Each observer sweeps the classroom by starting in one corner and recording each child's behavior in turn. In the KEEP school, the search and sweep methods have yielded equivalent overall means.

A three-category system is presented here: on-task, off-task, and disruptiveness. Off-task includes not being appropriately engaged, but does not include disruptiveness, which is defined as any behavior which interferes with or distracts from the work of other children or the teacher.

The comparison schools reported below are mostly the same schools included in the teacher data; some, although not all, observations were of the same sessions.

Comparison data will be presented for three schools observed in 1973-74 and for one school observed in 1975 (see Table 4). For both years, the problem of observer scheduling makes comparisons inexact in terms of time. While research assistants collected data in the field, often no one was available to collect KEEP data the same day. In such cases, for purposes of comparison, the nearest day preceding or following the public schools observation was chosen. All data were collected in the morning during reading and/or language arts. With the exception of kindergarten readiness in the fall at School 3, only KEEP data comparable in time and subject were used. All comparisons are for equivalent

Table 4

Summary On-Task, Off-Task, Disruptive Means and Standard Deviation: KEEP and Comparison

Group	Level	No. of Obser.	On-Task $\bar{X}$	On-Task SD	Off-Task $\bar{X}$	Off-Task SD	Disruptive $\bar{X}$	Disruptive SD
KEEP	K-1	3	84.00*	10.39	16.00	10.39	0	0
School 1	K-1	3	60.33	7.37	34.33	1.52	5.33	8.39
KEEP	K-1	4	91.75**	9.25	8.25	9.25	0	0
School 2	K-1	4	74.25	4.64	23.25	5.90	2.50	2.38
KEEP	K-1	9	85.11	9.04	14.44	8.27	.66	.41
School 3	K	17	82.17	7.35	14.23	7.26	3.64	2.20
KEEP	K-3	52	87.90*	5.83	11.64	5.18	.13	.25
School 5	2-3	6	73.83	7.68	26.16	7.68	0	0
KEEP	2-3	33	84.18*	8.82	16.10	7.35	.06	.24

* $p < .05$

** $p < .025$

grade level, and type of activity. Observer reliability obtained at KEEP in 1973-74 was 81% for categories combined, across all pairs of observers; the reliability for on-task itself ranged from .82 to .85. In 1975-76, overall reliability was .99.

School 1. Fifty-eight kindergarten and first grade children were observed. They were in a Hawaii English Program class, held in a single classroom and taught by three female teachers.

The teachers were to be participants in a workshop in positive reinforcement held at KEEP in the spring of 1974. The data presented are for pre-workshop "baseline" data. The data were collected for three days late in April (see Technical Report #33).

Table 4 presents the means and standard deviations for the three observed behaviors for KEEP and for School 1. An independent t-test showed that the rate of on-task was significantly higher at KEEP (t (4)=3.21, $p < .05$).

School 2. Forty-six children in a combined kindergarten-first grade were observed. They were taught by three female teachers in a single classroom. (These data were collected as baseline measures for the same workshop as for School 1.) An independent t-test showed again that KEEP had a significantly higher on-task rate (t (6)=3.38, $p < .05$). (See Table 4)

School 3. Two self-contained kindergarten classrooms (Class A and Class B) at a school located in a public housing area from which many KEEP students come were observed. There were about 25 students in each class and each was taught by one female teacher. These data were collected in 1973-74 for the sole purpose of comparison with KEEP. They were collected in the morning over three days in the fall (kindergarten readiness), three days in the winter (two days for Class B), and three days in the spring.

KEEP had slightly higher (a mean of about 2% for both classes) on-task means than both classes of School 3 and higher off-task. Both classes of

School 3 have higher mean disruptive rates (see Table 4). None of these differences was statistically significant. Since the on-task rate is high for both KEEP and the comparison group, the differences also would seem to have little practical significance.

School 5.[4] The research technique described in the preceding section on coding procedures at KEEP in 1973-74 was used with one modification. This was: The observations were confined to one cycle rather than three (so that the observer could collect data on all classes at KEEP: about 105 students). The observations done in the field followed the procedure of previous years. The observer conducted sweep cycles at 15 minute intervals. Each sweep took about 10-15 minutes.

School 5 is a public elementary school which draws its pupils from a population similar to KEEP's. Three teachers jointly taught a combined classroom of 48-58 second and third grade children. Observations were taken in three days in November, using two to three cycles of sweeps per day.

Comparison data at KEEP were collected specifically for this report. In order to give a comprehensive picture, all teachers' classrooms were observed (Teachers=8). Thus the 52 observations, which averaged 88% on-task, are a fair representation of the overall KEEP school at that time.

Overall, KEEP had a higher on-task rate, about 88% compared to School 5's mean of about 74%. This difference was significant (t (7)=2.74, $p < .05$).[5]

We can conclude, then, that the experimental KEEP pupils are more industrious than most sister-schools' pupils: The most recent data indicate perhaps 20% more (see Table 4). That 20% makes for a very perceptible difference in ambience. During the morning, when reading instruction occurs, our children are indeed hard at it. Visitors invariably note this industriousness. "Humanistically" oriented colleagues often are disturbed by the "regimentation"; visiting mothers are pleased.

4. On-task data are not available for School 4.

5. See Technical Report #55 for manner of data reduction.

Since our goal was an externally valid test of social reinforcement techniques and all variables could not be controlled, it is possible that the results on industriousness presented can be attributed to causes other than teacher use of reinforcement, for example, changed attitudes. In view of the mass of published evidence on social reinforcement effects, however, we think it probable that there is a causal link between teacher praise and pupil on-task rate. At this point, the link is strictly inferential since, in a normal classroom, it is difficult to demonstrate the praise-on-task link as a strict reinforcement function; modeled reinforcement effects are present, curriculum content is often carried by the praise, and child behaviors do not occur with the purity which allows precise shaping. Further, requirements of design affect teachers and the classroom operation so that both lose external validity.

However, we believe the pattern of teacher behavior is the cause of the better on-task rate, and that our results add to the rising stack of experimental data arguing that positive social reinforcement by the teacher is a most effective management technique for the ordinary classroom. We are also persuaded by the fact that the reinforcement manipulation was the only major manipulation in these first KEEP years. One cannot argue that we have an inherently more interesting curriculum, because we kept it comparable to local public schools.

The Reading Curriculum

During the first, planning year of KEEP operation, our staff visited very widely in kindergarten-third grade classrooms throughout the state, and consulted extensively with administrators and central office personnel. Although Hawaii has a unified, state-wide system, its reading curriculum is in a state of flux, as is typical of any vigorous, concerned system. A number of curricula were then employed in the state, including one of local design then being newly and widely adopted, the Hawaii English Program or HEP (the subsequent results from

which have been disappointing). Thus, all of our "comparison" schools do not use the same curriculum but we chose the Ginn 360 series because it is of demonstrated effectiveness nationally, and it is widely used locally. It is a middle-of-the-road basal reader program, rather conservative, with an emphasis on phonics. The Ginn series, at KEEP, was enriched with a bit of language-experience teaching, and efforts to make reading meaningful through group activities, etc., but these variations were confined to the range typical of our sister-schools. Our effort was primarily to teach to the Ginn curriculum.

It must be emphasized that there was not then, nor is there now, any curriculum of demonstrated effectiveness with the population at question. While every visitor to our center has one or another pet curriculum, none have favorable data to report on its effects on Hawaiian children. Needless to say, there is no curriculum developed through a study of Hawaiian children. Our long-range intention is to do just this, and against that day we have collected a mass of linguistic and cognitive information about our pupils. We are now beginning to design that curriculum, but that anticipates our story.

The Ginn 360 was chosen because it is a good, sound, ordinary curriculum, and our children attacked it industriously.

Results in Learning to Read

The Ginn series lends itself to detailed record keeping. During 1973-75, each objective within the curriculum was identified, and detailed records were kept on the number of days of instruction required for the mastery of each objective for each child. These data allowed several subsidiary studies, but will not be reported here, because the number-of-objectives achieved correlates (second grade: $r=.92$; first grade: $r=.79$) with end-of-year scores on the standardized Gates-MacGinitie Reading Test. Gates-MacGinitie results are transformed into percentile scores, with 50 representing the mean established in the

standardization group, and thus a national "norm." Two subscales, Vocabulary and Comprehension, are derived, as well as a Composite, or overall, score.

At this date, data are available on two KEEP classes only, although KEEP Class I has been with us long enough to report their scores both at the ends of first and second grades. Class II data are for end of first grade testing.

Data for four comparison schools are also presented (see Table 5). The Rural School is largely populated by Hawaiian children and is located on a neighbor island. Its reading curriculum is the Hawaii English Program (HEP). The Urban Schools are near KEEP, and draw from the same population. (These are Schools 1 and 3, for which teacher behavior and on-task rates are reported above. School 1 uses Distar; School 3, HEP.) The Suburban School is populated by middle-class Oriental and Caucasian children who typically prosper in the public schools. It also uses the Ginn 360 curriculum.

While KEEP scores may be gradually increasing, it can be seen that our children's achievement in reading does not differ significantly from other schools' Hawaiian children.

It will be recalled that KEEP's on-task rate was not notably superior to School 3, but it was superior to School 1. Yet among these three and the Rural Hawaiian School (using three different curricula), there are no real differences in reading. Even if the data could be tortured into statistical significance, there is certainly no social significance among programs which hover at $\leq$ 15th percentile.

These Gates-MacGinitie scores cannot be dismissed as inappropriate measures. In the first place, the Suburban School's scores are consistent with its overall record of producing students somewhat above national averages. It uses the same curriculum as does KEEP, and, in the most recent statewide testing of fourth grade students on the Stanford Achievement Test, it is clear that the

Table 5

Gates Mean Standard Score, Percentile, & Grade Score: KEEP & Comparisons

Group	Vocabulary			Comprehension			Composite	
	S. S.	%tile	Grade Score	S. S.	%tile	Grade Score	S. S.	%tile
KEEP I*	34.50	6	1.3-1.4	37.25	10	1.4	35.54	7-8
Rural *	40.08	16	1.4	40.50	16	1.5	40.29	16
Urban 1*	33.50	4.5	1.3	37.91	12	1.5	35.70	7-8
Suburban*	51.71	58	2.3	50.60	53	2.0	51.14	54
KEEP I**	38.08	13	1.6	35.29	7.5	1.6	36.71	8-10
Rural 2d**	43.55	25.5	2.3-2.4	39.67	15	1.9-2.1	41.33	18-21
Urban 1 2d**	33.86	4.5	1.4	37.81	11	1.7	35.83	7-8
Urban 2 2d**	42.74	22.5	2.2	38.22	13	1.8	40.48	16-18
Suburban 2d**	53.82	62	3.5-3.6	51.18	54	3.1-3.4	52.50	58-62
KEEP II**	41.71	19.5	1.6	37.82	11.5	1.4-1.5	39.61	14-16
Rural 1st**	40.53	17	1.5-1.6	33.67	4.5	1.3	35.77	10-12
Urban 2 1st**	39.38	15	1.5	35.94	7.5	1.4	37.97	10-12
Suburban 1st**	52.58	60	2.3-2.4	53.81	64	2.1-2.2	53.18	62-66

* 1973-74 administration (1st grade only)

** 1974-75 administration

modal score achieved by schools with predominantly Hawaiian children is Stanine One (sic), (Honolulu Advertiser, January 28, 1976).

Yet, the difficulty KEEP children suffer is relatively specific to reading. For example, while scoring in the 15th percentile on reading, KEEP's Class II had a Metropolitan Total Math score at the 40th percentile.

In Mathematics instruction, we use a standard curriculum, the Addison-Wesley, which, coupled with high on-task rate, produces a respectable achievement, until the children encounter word problems, at which time reading comprehension difficulties plummet their scores. For example, our Class I, by the end of second grade, reaches only the 26th percentile on the Math test.

Further, one cannot attribute low reading achievement to a low general intelligence in this population. Table 6 presents WPPSI means for KEEP children obtained at pre-enrollment in kindergarten (Fall), and during the last month of kindergarten (Spring).

These children do have a higher Performance than Verbal IQ, but after one year at KEEP they are within the normal range of intelligence, and their scores are hardly commensurate with reading at the 15th percentile, particularly when they work with notable industriousness.

Discussion

In spite of a successful program of teacher reinforcement, and associated solid industriousness by the pupils, KEEP children have achieved a dismal level of learning to read, when a standard curriculum is employed.

While these results do not argue against the use of reinforcement principles (which apparently do result in industrious pupils), they caution against over-reliance on them. Our interpretation of these findings is that the curriculum is inappropriate. When a child of normal intelligence performs as he is instructed and does not learn, the instructions are inadequate.

Table 6

WPPSI Mean IQs Fall Kindergarten

and Spring Kindergarten: Classes I, II, and III

	Total I.Q. Full Scale		Verbal I.Q.		Performance I.Q.	
	Fall	Spring	Fall	Spring	Fall	Spring
Class I 1972-73 N=27	90.19	101.31	86.22	92.70	97.12	110.93
Class II 1973-74 N=27	83.81	97.15	80.63	90.30	89.70	104.93
Class III 1974-75 N=28	91.32	105.21	87.00	100.32	97.79	109.46

The utility of reinforcement in the teaching of reading has been documented, for example by Staats and his associates (e.g., Staats, 1968; Staats and Butterfield, 1965; Staats, Minke, Finley, Wolf, and Brooks, 1964). It should be remembered, however, that Staats equally emphasizes the necessary step-by-step analysis of the component skills of the reading process. This analysis, and the programming of skill development through specific acts, results in curriculum.

The reading curricula available for Hawaiian children--not only the Ginn series, but all the others--have resulted from years of experience with a different population. By extension, this condition also obtains for other minority groups. Perhaps our results should lead us away from strategies advocated for minority group reading instruction which have as their goal an indirect increase in industriousness, whether the advocated route is through community control, or an increase in self-concept, or a more entertaining, "free" school. Desirable as any of these may be for important non-reading goals such as cultural heritage instruction, to the extent that they rely on a vigorous engagement of the child with the learning task, they will not suffice to teach reading. Teacher reinforcement has produced for us a relatively rapid and inexpensive means of directly achieving it, but industriousness is not enough.

The development of an appropriate curriculum, growing from a careful study of Hawaiian children, is now underway, and will be reported elsewhere.

References

Campbell, D. T. The social scientist as methodological servant of the experimenting society. Policy Studies Journal, 1973, 2 (1), 72-75.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30 (3), 116-127.

Gallimore, R., Boggs, J., & Jordan, C. Culture, behavior, and education. Beverley Hills: Sage Publications, 1974.

Johnson, J., & Sloat, K. C. M. The incremental effectiveness of guided practice and feedback in a course of behavior modification for teachers (Tech. Rep. #52). Honolulu: The Kamehameha Schools, Kamehameha Early Education Program, 1976.

Lam, D., Kidoguchi, L., Gallimore, R., Tharp, R. G., & Speidel, G. E. The uses and limits of increasing student motivation (Tech. Rep. #6). Honolulu: The Kamehameha Schools, Kamehameha Early Education Program, 1975.

MacDonald, W. S., & Gallimore, R. Battle in the classroom. Scranton, Pa.: Intext, 1971.

Sloat, K. C. M., Tharp, R. G., & Gallimore, R. The incremental effectiveness of classroom-based teacher-training techniques (Tech. Rep. #41). Honolulu: The Kamehameha Schools, Kamehameha Early Education Program, 1976.

Speidel, G. E., & Tharp, R. G. Training teachers in the use of positive feedback (Tech. Rep. #33). Honolulu: The Kamehameha Schools, Kamehameha Early Education Program, 1975.

Staats, A. W. Learning, language, and cognition. New York: Holt, Rinehart, & Winston, 1968.

Staats, A. W., & Butterfield, W. H. Treatment of non-reading in a culturally deprived juvenile delinquent: An application of reinforcement principles. Child Development, 1964, 3, 925-942.

Staats, A. W., Minke, K. A., Finley, J. R., Wolf, M. M., & Brooks, L. O. A reinforcer system and experimental procedure for the laboratory study of reading acquisition. Child Development, 1964b, 35, 209-231.

White, M. A. Natural rates of teacher approval and disapproval in the classroom. Journal of Applied Behavior Analysis, 1975, 8 (4), 367-372.